VOLUME 26 NUMBER 4

ACCOUNTS OF CHEMICAL RESEARCH®

APRIL 1993



# Methodological Advances in Protein NMR

AD BAX* AND STEPHAN GRZESIEK*

*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892*

*Received August 13, 1992*

Since the first experimental observation of nuclear magnetic resonance (NMR) in bulk matter more than 45 years ago,[1,2] its history has been punctuated by a series of revolutionary advances that have greatly expanded its horizons. Indeed, methodological and instrumental developments witnessed over the past two decades have turned NMR into the most diverse spectroscopic tool currently available. Applications vary from exploration of natural resources[3] and medical imaging to determination of the three-dimensional structure of biologically important macromolecules.[4-9] The present Account focuses primarily on the methodological advances in this latter application, particularly as they relate to the study of proteins in solution.

After Ernst and Anderson developed Fourier transform NMR[11] and demonstrated its use, the introduction of a second frequency dimension in NMR spectroscopy by Jeener in 1971[12] provided a critical trigger to the development of this field. An enormous variety of experimental schemes, all based on the two-dimensional (2D) concept and largely developed by the group of Ernst,[13] expanded the applicability of NMR to the characterization of quite complex molecules, including natural products, sugars, synthetic polymers, and peptides. Nearly a decade ago, protein structure determination was added to the realm of applications by the introduction of new systematic procedures for spectral analysis, primarily developed by Wüthrich and co-workers.[4] During the 1980s, development of new experimental pulse schemes continued to increase the power and applicability of 2D NMR to structural characterization of biopolymers. The most important development was undoubtedly the addition of a third frequency dimension to the NMR spectra.[14,15] The concept of 3D NMR is so similar to 2D NMR that no new formalism for the description of such experiments is required. The main problem that had to be solved for the development of such techniques was a way to record and process the enormous data matrices associated with such experiments. The second problem, as will be discussed later, was that sensitivity of the 3D experiments frequently is much lower than for analogous 2D experiments unless the third dimension corresponds to the chemical shift of a $^{13}C$ or $^{15}N$ nucleus and isotopic enrichment is used.[16,17] The advances in genetic engineering techniques that have occurred in the last decade enable many proteins to be overproduced

Ad Bax was born in the Netherlands. In 1981, he received his Ph.D in applied physics from the Delft University of Technology after conducting most of his graduate research with Ray Freeman in the Physical Chemistry Laboratory at Oxford University, working on the development of two-dimensional NMR methods. Following a postdoctoral appointment at Colorado State University, working in solid-state NMR, he joined the National Institutes of Health, where he presently is Chief of the Section on Biophysical NMR Spectroscopy. His work focuses on the development of improved NMR methods for the characterization of the structure and dynamics of biological macromolecules.

Stephan Grzesiek was born in West Germany and received his Ph.D in physics from the Free University of Berlin in 1988 for optical studies of proton release in bacteriorhodopsin. After a venture in a small software company, he received a fellowship from the Roche Research Foundation for NMR studies of pharmaceutically relevant biomolecules. Part of his fellowship he spent at the NIH, where he presently is employed as a Visiting Associate in the Section on Biophysical NMR Spectroscopy.

(1) Bloch, F.; Hansen, W. W.; Packard, M. *Phys. Rev.* **1946**, *69*, 127.
(2) Purcell, E. M.; Torrey, H. C.; Pound, R. V. *Phys. Rev.* **1946**, *69*, 37-38.
(3) Jackson, J. A. *Log Analyst* **1984**, 16-30.
(4) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(5) Clore, G. M.; Gronenborn, A. M. *Crit. Rev. Biochem. Mol. Biol.* **1989**, *24*, 479-564.
(6) Kessler, H.; Gehrke, M.; Griesinger, C. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 490-536.
(7) Bax, A. *Annu. Rev. Biochem.* **1989**, *58*, 223-256.
(8) Wüthrich, K. *Acc. Chem. Res.* **1989**, *22*, 36-44.
(9) Kaptein, R.; Boelens, R.; Scheek, R.; van Gunsteren, W. F. *Biochemistry* **1988**, *27*, 5389-5395.
(10) Clore, G. M.; Gronenborn, A. M. *Science*, **1991**, *252*, 1390-1399.
(11) Ernst, R. R. Adv. Magn. Reson. **1966**, *2*, 1-137.
(12) Jeener, J. Ampere Summer School, Basko Polje, Yugoslavia, 1971.

and labeled in microorganisms with the NMR observable stable isotopes. With isotopic enrichment, sensitivity of many of the heteronuclear 3D experiments is sufficiently high to add yet another frequency dimension to the NMR spectrum, dispersing resonance frequencies in four orthogonal dimensions.[18-20] In this Account, we will discuss the advantages and problems associated with extending the dimensionality of the NMR spectrum and will attempt to provide an answer to the question, How many dimensions do we really need?

## I. Principles of Multidimensional NMR

Although the principles of 2D NMR have been reviewed many times, we will briefly reiterate some of these in order to clarify the basis of 3D and 4D NMR. Most of the useful *n*D NMR experiments are of the so-called "correlated" type, in which the chemical shift of a nucleus is correlated with the chemical shifts of other nuclei based on an interaction between them. For example, in the important 2D NOESY experiment, protons are correlated on the basis of the dipole–dipole coupling between their magnetic moments, giving rise to magnetization transfer via the nuclear Overhauser effect (NOE). The pulse scheme used for the NOESY experiment is sketched in Figure 1a. In this scheme, three RF pulses are applied to the proton spins, and the scheme is repeated many times for systematically incremented durations of the time $t_1$. The signals detected during the time $t_2$ are modulated by the frequencies present during the time $t_1$, and a two-dimensional Fourier transformation of the acquired data matrix results in the 2D NMR spectrum.

In the NOESY spectrum, correlations between a resonance of proton A and proton B will be observed if A and B are sufficiently close in space (less than ~5 Å). However, before the distance information in the NOESY spectrum can be fully interpreted, it is necessary to assign each of the resonances in the $^1$H NMR spectrum to its site in the chemical structure. To accomplish this it is also necessary to record so-called *J*-correlated experiments, in which magnetization is transferred between chemically bonded nuclei via the *J*-coupling mechanism. The oldest 2D NMR pulse scheme, in which magnetization is transferred from one proton to another via $^1$H–$^1$H *J* coupling, is known as the COSY experiment and is probably the most popular experiment in the NMR analysis of small molecules. This experiment requires that $J_{HH}$ be not much smaller than the $^1$H resonance line width. This line width is approximately proportional to the inverse of the molecular tumbling rate and therefore increases approximately linearly with the size of the protein. For larger proteins $^1$H–$^1$H *J* couplings are frequently smaller than the line width, making the COSY experiment ineffective. Other *J* correlation techniques, such as the one depicted in Figure 1b, can correlate the frequency of a proton with that of its directly attached heteroatom ($^{13}$C or $^{15}$N). The heteronuclear one-bond couplings, $^1J_{CH}$ (125–160 Hz) and $^1J_{NH}$ (~92 Hz), are much larger than $^3J_{HH}$, and frequently as much as 50–90% of the magnetization can be transferred from protons to their directly coupled heteronuclei.[16] Consequently, such 2D heteronuclear shift correlation techniques are highly sensitive and can often be carried out even without isotopic enrichment.

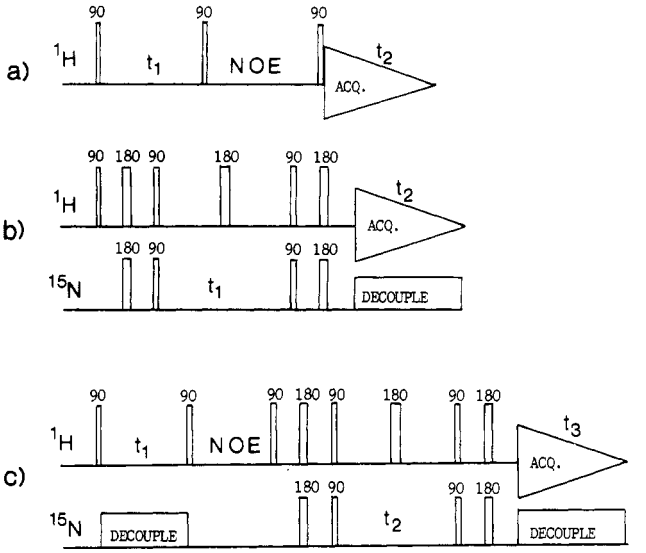


**Figure 1.** Examples of timing diagrams of 2D and 3D NMR pulse sequences. (a) 2D NOESY experiment. (b) 2D $^1$H-detected $^1$H–$^{15}$N HSQC correlation experiment. (c) 3D pulse scheme for 3D $^{15}$N-separated NOESY-HSQC experiment, obtained by concatenating schemes a and b. Radiofrequency pulses are marked by vertical bars and have typical durations of tens of microseconds. Signal is acquired during the time $t_2$ (schemes a and b) and $t_3$ (c); each scheme is repeated many times while the duration of $t_1$ (and $t_2$, for scheme c) is systematically incremented from 0 to ~30 ms.

The concept of 2D NMR is easily extended to higher dimensionality. For example, the two pulse schemes of Figure 1a,b can be concatenated in a manner depicted in Figure 1c, yielding a 3D experiment. The signals, acquired during the time $t_3$, are now obtained for many different $t_1$ and $t_2$ durations. As was the case in the 2D NOESY experiment, the data are modulated in the $t_1$ dimension by the frequencies of other nearby protons; however, in the $t_2$ dimension the modulation frequency is that of the $^{15}$N nucleus that is directly attached to the observed proton. Consequently, a 3D Fourier transformation (with respect to the time variables $t_1$, $t_2$, and $t_3$) yields a 3D frequency domain NMR spectrum. For practical reasons, such a spectrum is usually displayed and analyzed as a series of adjacent cross sections through the 3D data matrix.

Figure 2 illustrates the practical advantage of 3D over 2D NMR. The region of the conventional 3D spectrum that displays the NOE interactions involving amide protons of the protein staphylococcal nuclease (156 residues) is shown in Figure 2b and exhibits a high degree of resonance overlap. Figure 2a represents a

(13) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Clarendon Press: Oxford, 1987.
(14) Vuister, G. W.; Boelens, R.; Kaptein, R. *J. Magn. Reson.* **1988**, *80*, 176–185.
(15) Oschkinat, H.; Griesinger, C.; Kraulis, P. J.; Sørensen, O. W.; Ernst, R. R.; Gronenborn, A. M.; Clore, G. M. *Nature (London)* **1988**, *332*, 374–376.
(16) Fesik, S. W.; Zuiderweg, E. R. P. *J. Magn. Reson.* **1988**, *78*, 588–593.
(17) Marion, D.; Kay, L. E.; Sparks, S. W.; Torchia, D. A.; Bax, A. *J. Am. Chem. Soc.* **1989**, *111*, 1515.
(18) Kay, L. E.; Clore, G. M.; Bax, A.; Gronenborn, A. M. *Science* **1990**, *249*, 411–414.
(19) Clore, G. M.; Kay, L. E.; Bax, A.; Gronenborn, A. M. *Biochemistry* **1991**, *30*, 12–18.
(20) Zuiderweg, E. R. P.; Petros, A. M.; Fesik, S. W.; Olejniczak, E. T. *J. Am. Chem. Soc.* **1991**, *113*, 370–372.

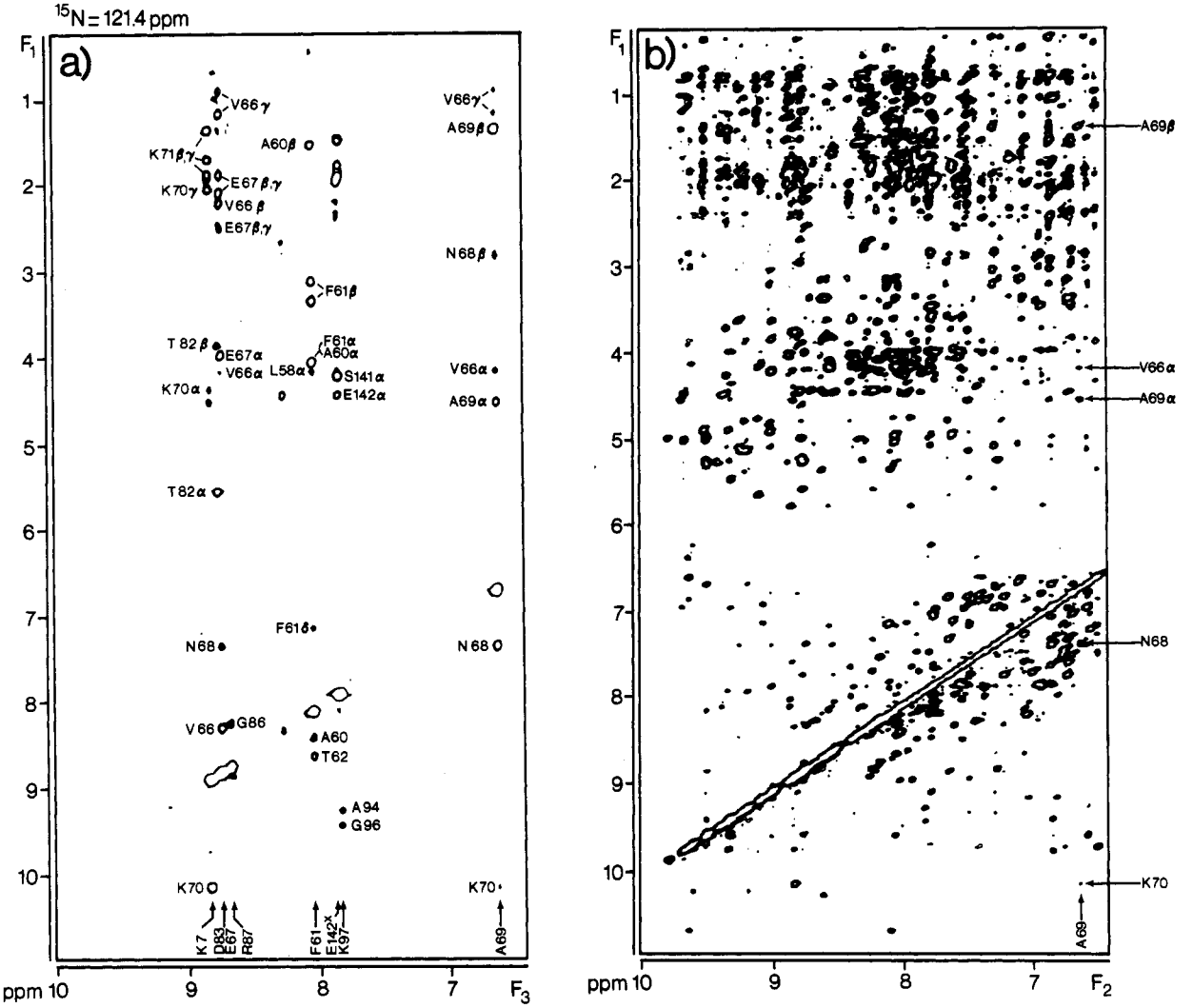


**Figure 2.** (a) One out of 64 parallel cross sections through the 3D $^{15}$N-separated NOESY spectrum of the protein staphylococcal nuclease, displaying NOE interactions involving amide protons that are attached to a $^{15}$N with a 121.4 ± 0.3 ppm chemical shift. (b) Corresponding region of a regular 2D NOESY spectrum, displaying NOE interactions involving all amide protons. From ref 17.

slice taken through the 3D $^{15}$N-separated NOESY spectrum and displays NOE interactions only for amide protons attached to $^{15}$N nuclei with a chemical shift near 121.4 ppm. The entire 3D spectrum consists of 64 such slices, representing interactions to amides with $^{15}$N chemical shifts ranging from 130 to 105 ppm.

Although resonance overlap in the 3D spectrum is dramatically reduced compared to that in 2D, interpretation of the 3D NMR spectrum frequently is not necessarily straightforward. First, even if the chemical shift frequency is known for each proton in the protein, these shifts are often insufficiently unique to identify the proton. For example, if in the 3D NOESY spectrum we observe that the amide proton (at 6.7 ppm) of Ala-69 ($^{15}$N at 121.4 ppm) interacts with a proton at 4.2 ppm, this does not identify uniquely the second proton because nearly a dozen protons resonate in the 4.2 ± 0.02 ppm region. Therefore, it is useful to disperse the 3D spectrum in yet another dimension in order to reveal the frequency of the $^{13}$C directly attached to the proton at 4.2 ppm. The pair of $^{1}$H and $^{13}$C shifts associated with the second proton frequently identifies it in a unique or nearly unique manner, greatly facilitating the identification of NOE interactions. Extending the 3D experiment into four dimensions is straightforward and involves inserting a 2D $^{1}$H–$^{13}$C correlation scheme in the 3D experiment of Figure 1c.[18] The power of 4D NMR is illustrated in Figure 3, for the amide proton of Asn-17 in the protein interferon-γ, a homodimer with a total molecular weight of 31.4 kDa. Figure 3a shows a "strip" taken through the 3D spectrum at the $F_2$ and $F_3$ coordinates of the Asn-17 amide, and it shows the chemical shifts of all protons that have an NOE interaction with this amide proton. Figure 3b shows the corresponding slice through the 4D spectrum, and it yields both the $^{1}$H and $^{13}$C chemical shifts.

Clearly, 4D NMR presents a desirable and logical approach for identification of NOE interactions in larger proteins. The method is more "symmetrical" than the 3D $^{15}$N- or $^{13}$C-separated NOESY experiments, where the chemical shift of its attached heteronucleus is obtained for only one of the two interacting protons. The 4D method mentioned above is only suitable for separation of NOEs between $^{15}$N- and $^{13}$C-attached protons. A conceptually similar, but technically more demanding, 4D experiment can separate NOE interactions between $^{13}$C-attached protons.[19,20] This is crucial for the unambiguous identification of interresidue side chain–side chain NOE contacts.

**Assignment Approach.** For small proteins, assignment of the $^{1}$H spectrum can be made by combined analysis of NOE- and $^{1}$H–$^{1}$H $J$-correlated 2D spectra.

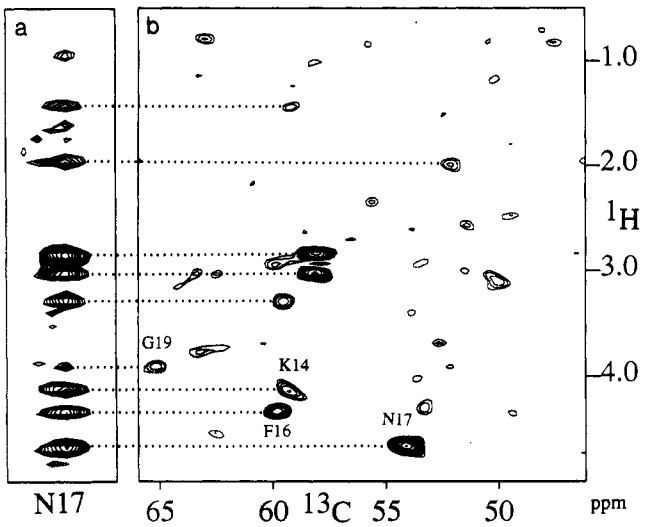


**Figure 3.** (a) Strip from the $^{15}$N-separated 3D NOESY spectrum of interferon-γ displaying the chemical shifts of aliphatic protons that have an NOE with the backbone amide of Asn-17. The strip is actually a narrow vertical band of a 2D cross section, such as shown in Figure 2a. (b) Cross section through the 4D $^{15}$N/$^{13}$C-separated NOESY spectrum, displaying the chemical shifts of the protons that have an NOE interaction to the amide proton of Asn-17, together with the shifts of the $^{13}$C nuclei directly attached to these protons. Broken contours correspond to $^{13}$C nuclei in the 46–26 ppm chemical shift range, which have been aliased once in the $^{13}$C dimension. Adapted from ref 48.

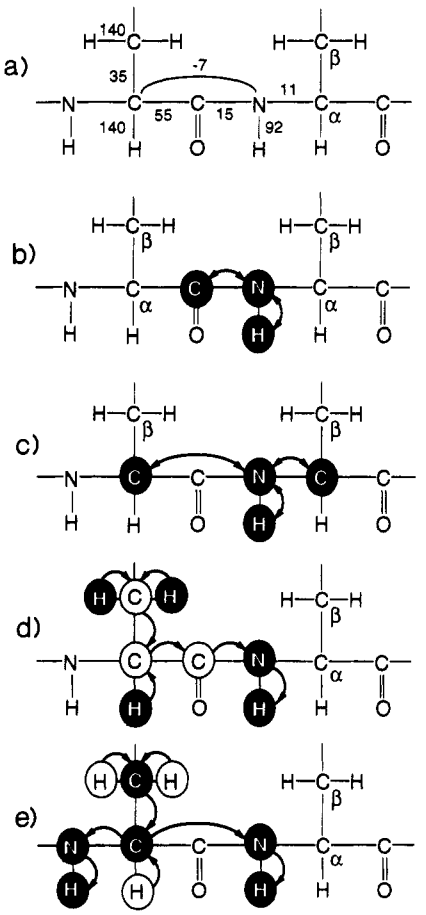


**Figure 4.** (a) A dipeptide segment of a protein backbone with the approximate values for the $J$ couplings which are essential for the assignment procedure in isotopically enriched proteins. (b–e) Schematic diagrams of the nuclei that are correlated in the (b) HNCO, (c) HNCA, (d) HBHA(CBCACO)NH, and (e) CBCANH experiments. Nuclei for which the chemical shift is measured in the 3D experiment are marked by solid circles. Nuclei involved in the magnetization transfer pathway, but not observed, are marked by open circles. Magnetization transfer in these experiments is marked by curved solid lines, and the direction of the transfer is marked by arrows.

However, as indicated earlier, $^1$H–$^1$H $J$ correlation techniques frequently fail for larger proteins because of the increased $^1$H resonance line width. This problem is even worse for proteins enriched with $^{13}$C, because the $^{13}$C–$^1$H dipolar interaction causes additional proton line broadening. As assignments cannot be made on the basis of NOE interactions alone, other experiments for obtaining through-bond $J$ correlations are essential. Moreover, in order to utilize the additional $^{15}$N and $^{13}$C chemical shift information available from the 3D and 4D $^{15}$N- and $^{13}$C-separated NOESY spectra, it is necessary to assign all protonated $^{15}$N and $^{13}$C nuclei.

In recent years, a novel assignment procedure has been developed that is applicable to uniformly isotopically enriched proteins.[21–24] This procedure is quite different from the traditional approach and is based primarily on one-bond $J$ couplings between adjacent atoms. The one-bond $J$ couplings are relatively uniform and depend only weakly on conformation. Typical values for the relevant coupling constants are indicated in Figure 4. Equally important are the magnitudes of the transverse relaxation times, $T_2$, which determine the resonance line widths. $T_2$ depends approximately linearly on the molecular tumbling rate, i.e., on molecular size, and inversely on the viscosity, but also on the degree of internal mobility and on local conformation. As a rough guide, residues that do not have a high degree of internal mobility in a globular protein of 20 kDa at 35 °C have line widths of ~12 Hz for the amide proton, ~7 Hz for the amide nitrogen in the $^1$H-coupled mode and ~4 Hz in the $^1$H-decoupled mode, ~15 Hz for $^{13}C_\alpha$, and ~25 Hz for a $^{13}$C-attached $^1H_\alpha$. The line width of the carbonyl carbon is dominated by chemical shift anisotropy and therefore proportional to the square of the applied magnetic field; for the 20-kDa protein, values of ~6 Hz are observed at 500-MHz $^1$H frequency. Comparison of these numbers with the $J$ values shown in Figure 4 indicates that for a protein of 20 kDa most one-bond $J$ couplings are significantly larger than the line widths. This means that magnetization can be transferred with high efficiency from one nucleus to its directly coupled neighbor. In this way a number of 3D $J$-correlated NMR experiments have been constructed, correlating the backbone and side-chain resonances in a manner schematically indicated in Figure 4. For example, the HBHA(CBCACO)NH experiment[28] correlates the amide $^1$H and $^{15}$N

(21) Oh, B. H.; Westler, W. M.; Derba, P.; Markley, J. L. *Science* **1988**, *240*, 908–911.
(22) Ikura, M.; Kay, L. E.; Bax, A. *Biochemistry* **1990**, *29*, 4659–4667.
(23) Bax, A.; Ikura, M.; Kay, L. E.; Barbato, G.; Spera, S. *Protein Conformations*; Ciba Foundation Symposium 161; Wiley: New York, 1991; pp 108–135.
(24) Clore, G. M.; Gronenborn, A. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1991**, *23*, 43–92.
(25) Marion, D.; Driscoll, P. C.; Kay, L. E.; Wingfield, P. T.; Bax, A.; Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1989**, *28*, 6150–6156.
(26) Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *96*, 432–440.
(27) Palmer, A. G., III; Fairbrother, W. J.; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Biomol. NMR* **1992**, *2*, 103–108.
(28) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1992**, *114*, 6291–6293.
(29) Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *99*, 201–207.
(30) Clubb, R. T.; Thanabal, V.; Wagner, G. *J. Magn. Reson.* **1992**, *97*, 213–217.

**Table I. Experiments for Determining Backbone Assignments and Secondary Structure of Isotopically Labeled Proteins**

| experiment | purpose | solvent | labeling[c] $^{15}N$ | labeling[c] $^{13}C$ | S/N | time,[a] days | ref |
|---|---|---|---|---|---|---|---|
| $^{15}N$-separated HOHAHA | $H_N(i)$, $N(i)$, $H_\alpha(i)$ | $H_2O$ | + | – | ± | 3 | 25 |
| $^{15}N$-separated NOESY | short range NOE | $H_2O$ | + | – | ● | 3 | 16, 17 |
| HNCO | $H_N(i)$, $N(i)$, $CO(i-1)$ | $H_2O$ | + | + | ++ | 2 | 22, 26 |
| HNCA[b] | $H_N(i)$, $N(i)$, $C_\alpha(i)/C_\alpha(i-1)$ | $H_2O$ | + | + | + | 2 | 22, 26 |
| HCACO[b] | $H_\alpha(i)$, $C_\alpha(i)$, $CO(i)$ | $D_2O$ | +/– | + | ++ | 2 | 22, 27 |
| CBCA(CO)NH | $C_\beta(i-1)/C_\alpha(i-1)$, $N(i)$, $H_N(i)$ | $H_2O$ | + | + | + | 2.5 | 28 |
| HBHA(CO)NH | $H_\beta(i-1)/H_\alpha(i-1)$, $N(i)$, $H_N(i)$ | $H_2O$ | + | + | + | 2.5 | 28 |
| CBCANH | $C_\beta(i)/C_\alpha(i)/C_\beta(i-1)/C_\alpha(i-1)$, $N(i)$, $H_N(i)$ | $H_2O$ | + | + | ± | 2.5 | 29 |
| HN(CA)CO[b] | $N(i)$, $H_N(i)$, $CO(i)$ | $H_2O$ | + | + | ● | 2.5 | 30 |

[a] Minimum measuring time, determined by the required digital resolution and the minimum number of phase-cycling steps. "++", "+", and "±" refer to the inherent sensitivity of the experiment. Experiments labeled ++ and + may be shortened significantly by the use of pulsed field gradient methodology. [b] For proteins with favorable resonance dispersion and sensitivity, these three experiments may not be necessary. [c] "+" indicates essential labeling, "–" indicates that no labeling should be used, and "+/–" indicates that labeling is not relevant.

resonances of one residue with the $H_\alpha$ and $H_\beta$ chemical shifts of its preceding residue. Thus, for each amide two or three (in the case of nonequivalent $H_\beta$ protons) resonances are observed in the 3D HBHA(CBCACO)-NH spectrum, and their coordinates in the $F_1$, $F_2$, and $F_3$ dimensions of the 3D spectrum correspond to the $^1H_\alpha$ and $^1H_\beta$ ($F_1$), $^{15}N$ ($F_2$), and $^1H_N$ ($F_3$) chemical shift frequencies.

Table I lists the experiments that are needed to make the backbone $^1H$, $^{13}C$, and $^{15}N$ assignments and to determine the secondary structure. Also indicated in this table is the approximate measuring time needed for acquiring each of these spectra. This measuring time should be considered to be only a rough estimate for proteins in the 15–25-kDa range at concentrations of ca. 1 mM in a volume of ~0.4 mL. Longer measuring times may be needed for proteins that require high digital resolution because of particularly severe resonance overlap, for more dilute samples, or for proteins approaching the molecular size limit of a particular experiment, in which case the protein line width seriously degrades the sensitivity of the experiment. On the other hand, shorter measuring times could be afforded for spectra that are relatively well resolved and for sample concentrations significantly higher than 1 mM. In favorable cases, complete backbone assignments may be obtained using only a subset of the experiments listed in Table I. However, for proteins with substantial overlap in the $H_\alpha$–$C_\alpha$ correlation, as typically encountered for larger proteins rich in $\alpha$-helices, the entire arsenal, including additional experiments not listed here, may be required.

For determining the tertiary structure of a protein, additional experiments need to be carried out for obtaining side chain resonance assignments and for collecting the required NOE data. For obtaining a "high-resolution" structure, even more experiments are typically needed to measure the numerous homo- and heteronuclear $J$ couplings and to determine stereospecific assignments of nonequivalent methylene protons and methyl groups in valine and leucine residues. Consequently, approximately 45–90 days of measuring time is presently required to gather all spectral information needed for determining a high-resolution protein structure. For very soluble proteins (> ~2 mM) smaller than ~20 kDa, it may be anticipated that recent advances in pulsed field gradient methodology may shorten this minimum.[31-33] However, for experiments involving NOE or $J$-coupling measurements, which are usually limited by low signal-to-noise ratios, no dramatic improvement is expected. At present, analysis of the multitude of 3D and 4D spectra is even more time consuming than the data acquisition itself, particularly as software for completely automated analysis of the NMR spectra is not yet available. Development of such software, although conceptually quite straightforward, is frought with many practical problems that have not yet been solved in a general manner.

## II. Sensitivity of Multidimensional NMR

The sensitivity of 3D NMR experiments has been treated in a rigorous manner by Griesinger et al.[34] Here, we will briefly discuss some of the most critical factors. The sensitivity of multidimensional NMR experiments is determined primarily by the efficiency of the magnetization transfer steps. For example, if in the NOESY experiment, sketeched in Figure 1a, only 0.1% of the nuclear spin magnetization of proton A is transferred to proton B, this immediately results in a 1000-fold reduction compared to the resonance intensity of proton A in a conventional single-pulse one-dimensional spectrum. Another factor that affects the sensitivity of multidimensional NMR is the decay caused by relaxation during the evolution periods of the experiment. For example, if for the NOESY experiment mentioned above high resolution in the final 2D spectrum is critical, it will be necessary to use relatively long acquisition times in the $t_1$ dimension, i.e., the pulse scheme must be repeated for a range of $t_1$ durations that stretches from 0 to several times the transverse relaxation time, $T_2$, of the protons. Note that the experiments with long $t_1$ durations carry little signal and, therefore, decrease the signal-to-noise ratio obtainable per unit of time. Another important but frequently overlooked detail is the fact that each time the spectral dimensionality is increased, the sensitivity drops by $2^{1/2}$ because both the real and imaginary component of the signal must be sampled in separate experiments. Consequently, 4D experiments, for example, are in-

(31) Vuister, G. W.; Boelens, R.; Kaptein, R.; Hurd, R. E.; John, B.; van Zijl, P. C. M. *J. Am. Chem. Soc.* **1991**, *113*, 9688–9690.
(32) Davis, A. L.; Boelens, R.; Kaptein, R. *J. Biomol. NMR* **1992**, *2*, 395–400.
(33) Bax, A.; Pochapsky, S. S. *J. Magn. Reson.* **1992**, *99*, 638–643.
(34) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Magn. Reson.* **1989**, *84*, 14–63.

herently 2 times less sensitive than 2D experiments. An ingenious approach, applicable to a number of $J$-correlated experiments, which reduces this loss at the expense of an increase in the complexity of the experiment, has recently been proposed by Palmer et al.[35]

As indicated above, the efficiency of the magnetization transfer processes is the main factor determining the sensitivity. Consequently, many of the original homonuclear 3D experiments, which contain two relatively inefficient homonuclear magnetization transfer steps, require concentrated samples to overcome their low inherent sensitivity. Nevertheless, such 3D experiments can be extremely powerful in resolving ambiguities that are invariably present in 2D spectra of proteins in the 5–15-kDa molecular weight range.

**How Many Dimensions Do We Need?** With the recent introduction of a fourth dimension in NMR spectroscopy, the logical question to ask is what the optimal dimensionality of an NMR spectrum is. There is no single answer to this question, but the following discussion is intended to clarify some of the issues. As pointed out by Sørensen,[36] 3D or 4D NMR spectra can be considered as mathematical products of their corresponding 2D building blocks. Consequently, the 3D or 4D experiments do not offer fundamentally new information but merely resolve overlap problems present in the 2D spectra. Thus, the main purpose of increasing the dimensionality of the NMR spectrum is to reduce spectral overlap. Alternatively, the resolution in the 2D spectrum may be increased for the same purpose by using longer acquisition times in the orthogonal time dimensions. However, as indicated above, extending the acquisition time much beyond the applicable transverse relaxation time $T_2$ rapidly decreases the inherent sensitivity of the experiment, and the payoff in increased resolution associated with longer $t_1$ acquisition times drops sharply. Therefore, if acceptable spectral resolution cannot be obtained with $t_1$ acquisition times on the order of the applicable $T_2$, increasing the spectral dimensionality will be useful, provided that an efficient additional magnetization transfer step is available. For proteins isotopically enriched with $^{13}C$ and/or $^{15}N$, such an additional transfer usually can be generated quite efficiently.

Once the spectrum is adequately resolved, it usually does not pay to increase the spectral dimensionality any further. Consider, for example, a $J$-correlated experiment in which we are trying to correlate the $^{13}C_\alpha$ and $^1H_\alpha$ resonances of one residue with the $^{15}N$ and $^1H_N$ backbone amide resonances of the next residue, and suppose, for convenience, that the 2D $^{15}N$–$^1H$ correlation spectrum does not yield any spectral overlap. We now have the choice of correlating resonances of all four nuclei simultaneously in a 4D experiment or conducting two 3D experiments, one which correlates the amide resonances with $^{13}C_\alpha$, and one which correlates them with $^1H_\alpha$. To make the example less abstract, let us assume identical acquisition times in the $^{15}N$ and $^1H_N$ dimensions for the 3D and 4D experiments, and let us also assume that 8 complex increments are taken in the $^{13}C_\alpha$ and in the $^1H_\alpha$ dimension of the 4D experiment, i.e., the 2D $^{15}N$–$^1H$ correlation experiment must be repeated $(2 \times 8) \times (2 \times 8) = 256$ times. In the same amount of time, two 3D experiments could be recorded with 64 complex increments each in their respective $^1H_\alpha$ and $^{13}C_\alpha$ dimensions. In this case, the two 3D experiments determine the peak position of the $^1H_\alpha$/$^{13}C_\alpha$ pair up to 64 times more precisely than does the 4D experiment. This example merely serves to illustrate that the highest possible dimensionality is not necessarily always the best choice. On the other hand, as argued by Boucher et al.,[37] there may be other practical reasons that make it preferable to obtain information from a single 4D experiment instead of from two separate 3D spectra that have been acquired at different times, possibly under different conditions. Clearly, there are advantages and disadvantages associated with increasing the dimensionality of the NMR spectrum. In practice, an increase in the spectral dimensionality requires that an additional, efficient magnetization transfer step be available. For example, in the 4D NOESY experiment, the two transfer steps, added to the 2D NOESY scheme, involve correlating the two protons with their directly attached heteroatoms. With few exceptions, however, increasing the dimensionality in non-NOESY type experiments is necessary and beneficial only if resolution in the 3D spectrum is limited by the natural resonance line width, and not by the use of acquisition times that are shorter than the transverse relaxation time, $T_2$. If resolution is not limited by $T_2$, a combination of two or three experiments of the lower dimensionality frequently is more efficient at yielding the desired information than is an increase in dimensionality.

## III. Structural Parameters

To date, NMR protein structures have been calculated almost exclusively on the basis of interproton distances (derived from NOE measurements) and dihedral angles (derived from $^1H$–$^1H$ $J$ couplings). In addition, slow exchange of amide protons with water is usually interpreted as an indication of hydrogen bonding.

Over the past decade, complete $^1H$ resonance assignments have been made for a large number of proteins for which the 3D structure is accurately known from X-ray crystallographic studies. Careful analysis of the $^1H$ chemical shifts indicates that they can be predicted on the basis of the protein structure to within a few tenths of a part per million.[38] Therefore these chemical shifts may become useful indicators for the accuracy of a protein structure, or alternatively, they potentially could be used for further structure refinement.

The introduction of uniform isotopic enrichment yields access to additional structural parameters that may increase further the level of detail at which both the structure and the internal dynamics of a protein can be studied. These parameters are briefly discussed below.

**$^{15}N$ and $^{13}C$ Chemical Shifts.** The new isotope-assisted methodology (Table I) relies on $^{13}C$ and $^{15}N$ chemical shifts to resolve the very crowded $^1H$ spectra

(35) Palmer, A. G., III; Cavanagh, J.; Byrd, R. A.; Rance, M. *J. Magn. Reson.* **1992**, *96*, 416–424.
(36) Sørensen, O. W. *J. Magn. Reson.* **1990**, *89*, 210–216.
(37) Boucher, W.; Laue, E. D.; Campbell-Burk, S.; Domaille, P. J. *J. Am. Chem. Soc.* **1992**, *14*, 2262–2264.
(38) Ösapay, K.; Case, D. A. *J. Am. Chem. Soc.* **1991**, *113*, 9436–9444.

of larger proteins. The $^{13}C$ and $^{15}N$ resonances themselves also contain valuable information, however. For example, a clear correlation has been found between the protein backbone angles $\phi$ and $\psi$ and the $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts.[39,40] The $^{13}C_\beta$ resonance follows the same trend as the $H_\alpha$ proton resonance: for extended structures, with $\phi, \psi \sim 130°$, on average a downfield shift from the random coil position is observed, whereas for $\alpha$-helical structures ($\phi, \psi \sim -50°$), a small upfield shift is observed. The $^{13}C_\alpha$ resonance follows the opposite trend, with significant downfield shifts ($\sim$3 ppm) for helical structures and an upfield shift ($\sim$1–2 ppm) for extended structures.

The $^{15}N$ chemical shifts of backbone amides in proteins can also deviate substantially from their random coil values. For example, the $^{15}N$ shifts of Val-39 and Val-99 in staphylococcal nuclease differ by 32 ppm. Clearly, such deviations must be related to local structural differences. On average, amides in a $\beta$-sheet resonate downfield by about 5 ppm compared with $\alpha$-helical amides, but large exceptions to this rule occur. Attempts to correlate structural features with $^{15}N$ chemical shifts at a more detailed level so far have remained unsuccessful. Likely, there are a number of important contributors, including the $\phi$ and $\psi$ backbone angles, the planarity of the peptide bond, and hydrogen bonding of the amide proton and its adjacent carbonyl.

***J* Couplings.** To date, use of *J* couplings in protein structure determination has been restricted mainly to the three-bond $H_N$–$H_\alpha$ and $H_\alpha$–$H_\beta$ couplings. These couplings are correlated with dihedral angles following well-known Karplus equations and have been used to restrain the $\phi$ and $\chi_1$ angles, and for making stereospecific assignments of nonequivalent $H_\beta$ methylene resonances. In larger proteins, where these couplings typically cannot be obtained from the $^1H$–$^1H$ multiplet structures, they may be measured using a variety of different methods that utilize the presence of stable isotopes. These isotopes also provide access to a large number of heteronuclear *J* couplings, and the three-bond $^1H$–$^{15}N$ and $^1H$–$^{13}C$ couplings carry particularly important conformational information that can be interpreted readily using Karplus equations.

Although little used to date, one-bond $^1H$–$^{13}C$ *J* couplings are also related to structure in a simple manner, and these couplings can readily be measured in isotopically enriched proteins.[41] Quite recently, a technique has been described for measuring long-range $^{13}C$–$^{13}C$ *J* couplings in proteins.[42] For proteins in the 15–20-kDa range, these relatively small couplings ($< \sim 4$ Hz) can be measured only for methyl carbons which, because of their nonexponential transverse relaxation, have a narrow component to their resonance line shape. These *J* couplings provide very direct information on the side-chain torsion angles in Leu, Ile, Val, and Thr residues.

**$^{15}N$ and $^{13}C$ Relaxation Times.** The degree of internal protein flexibility not only is of fundamental interest but also is critical for understanding protein recognition. For proteins at natural isotopic abundance, the dynamic behavior of a protein is not easily quantified by NMR. In contrast, for proteins isotopically enriched with $^{15}N$ and/or $^{13}C$, this information can be retrieved from $^{15}N$ and $^{13}C$ longitudinal and transverse relaxation times and from the heteronuclear $^{15}N\{^1H\}$ and $^{13}C\{^1H\}$ NOEs.[43–45] The strength of this approach for the characterization of dynamics stems from the fact that the $^{15}N$ and $^{13}C$ relaxation times are dominated by the time dependence of the strong heteronuclear dipolar interaction with their attached proton(s). This time dependence is determined primarily by reorientation of the internuclear bond vector, and the relaxation measurements provide information on both the amplitudes of individual angular bond vector fluctuations and the time scale on which they occur. For practical reasons, most of the detailed protein relaxation measurements so far have focused on backbone amides. For example, we used this methodology to show that the so-called "central helix" of the protein calmodulin, which separates its two globular domains in the crystalline state, functions as a flexible linker in solution.[46] Specifically, the amide N–H bond vectors of four adjacent residues near the middle of this central helix show large angular fluctuations on a 100-ps time scale, and the anisotropy of the overall molecular tumbling, expected for a rigid "central helix", was not observed. In contrast, each of the globular domains reoriented in a nearly isotropic manner, with the time scale of the smaller of the two domains being faster than that of the larger domain.

Extending the relaxation measurements to aliphatic side chain carbons in principle is straightforward. However, in practice extra care needs to be paid to technical details, related to $^{13}C$–$^{13}C$ *J* couplings, which can affect measurements of $^{13}C$ transverse relaxation times and to analysis of the data when more than one proton is coupled to a particular $^{13}C$ nucleus. In the latter case, dipolar cross correlation can affect the quantitative interpretation of measured parameters. Although the theoretical formalism for describing this effect is well established, for proteins cross correlation can seriously affect the measured parameters themselves unless specific precautions are taken.[47]

## IV. Conclusion and Outlook

It is clear that the introduction of 3D and 4D NMR, combined with uniform $^{13}C$ and $^{15}N$ isotopic enrichment, significantly extends the molecular weight limit of proteins for which a solution structure can be determined by NMR. However, it is equally clear that such structural studies of larger proteins are time-consuming and costly. The minimum time for data acquisition needed for detailed characterization of a single protein in the 25–35-kDa range is expected to remain at least several months. Tens of milligrams of isotopically enriched protein, which must remain stable over such long measuring periods, are required for this process.

(39) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490–5492.
(40) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311–333.
(41) Vuister, G. W.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **1992**, *114*, 9674–9675.
(42) Bax, A.; Max, D.; Zax, D. *J. Am. Chem. Soc.* **1992**, *114*, 6924–6925.
(43) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4558.
(44) Dellwo, M. J.; Wand, A. J. *J. Am. Chem. Soc.* **1989**, *111*, 4571–4578.
(45) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972–8979.
(46) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269–5278.
(47) Kay, L. E.; Bull, T. E.; Nicholson, L. K.; Griesinger, C.; Schwalbe, H.; Bax, A.; Torchia, D. A. *J. Magn. Reson.* **1992**, *100*, 538–558.

Depending on the efficiency and the type of expression system, this can require large amounts of expensive isotopically enriched precursors. Analysis of the multitude of 3D and 4D NMR spectra, needed for a detailed structural characterization, presently also is a labor-intensive and time-consuming process. However, development of suitable support software is expected to alleviate this tedious burden in the near future, significantly reducing the time needed for structure determination.

The presently well established 2D methodology for determining protein structure starts to fail when the $^1$H line width, which is proportional to the protein molecular weight and inversely proportional to viscosity, becomes significantly larger than the homonuclear $^1$H–$^1$H $J$ couplings. For such larger proteins, resonance overlap in the 2D NOESY spectrum presents an additional serious barrier for structure determination. However, this latter limitation is less fundamental in nature since it may be resolved by going to higher field strengths or by recording homonuclear 3D experiments.[14,15]

For proteins enriched uniformly with $^{13}$C and $^{15}$N, the multinuclear methodology described in this Account works very well for intermediate size proteins but it starts to fail when the $^{13}$C line width becomes significantly larger than the $^{13}$C–$^{13}$C couplings. At room temperature this occurs for proteins of ~30 kDa. Indeed, for the protein interferon-$\gamma$ (31.4 kDa), line widths of ~45 Hz for the $^{13}C_\alpha$ resonances are observed at 27 °C, making complete assignments of all side-chain resonances extremely difficult.[48] Procedures for assignment of the backbone resonances are, in our experience, more robust and function quite well, even for proteins as large as interferon-$\gamma$. Recording of the 4D NOESY spectra also becomes problematic for these larger proteins because the sensitivity loss occurring during the $^1$H–$^{13}$C correlation step(s) in these experiments is significant. For example, in interferon-$\gamma$ the $^1$H line widths of many of the $^{13}$C attached protons are in excess of 50 Hz, resulting in a 3-fold loss in NOE cross peak sensitivity from relaxation during the $^1$H–$^{13}$C correlation step. Altogether, it therefore appears that 30 kDa presents an upper molecular weight limit for proteins that can be studied in detail with the technology outlined herein. Slightly larger proteins may be accessible if solvent viscosity can be decreased by raising the temperature or if the protein can be studied at concentrations significantly higher than 1 mM.

For completeness, it should be mentioned that the $^{13}$C- and $^{15}$N-based multidimensional approach described here is not the only possible approach for increasing the molecular weight limit of proteins that can be studied in detail by solution NMR. One powerful alternative approach utilizes either random or residue-specific deuteration, thus simplifying the $^1$H spectrum and narrowing $^1$H resonance line widths.[49,50] This approach has the advantage that the required NMR experiments are all relatively simple and of the homonuclear $^1$H type. However, when residue-specific labeling is called for, a large number of samples with different deuterated amino acid residues are needed, and obtaining a comprehensive set of resonance assignments remains a difficult and labor-intensive task. It is conceivable that partial $^2$H labeling in combination with $^{13}$C- and $^{15}$N-based multidimensional experiments may further increase the molecular weight limit of proteins whose solution structure can be determined by NMR. Alternatively, a combination of the $^{13}$C- and $^{15}$N-based multidimensional experiments and residue-specific $^{13}$C- and/or $^{15}$N-labeling may prove to be effective for this purpose.

*We thank Ted Becker, Marius Clore, Angela Gronenborn, David Live, John Schwab, Attila Szabo, and Dennis Torchia for stimulating discussions and useful suggestions during the preparation of this manuscript. The work in the authors' laboratory was supported in part by the Intramural AIDS Directed Anti-Viral Program of the Office of the Director of the National Institutes of Health.*

(48) Grzesiek, S.; Döbeli, H.; Gentz, R.; Garotta, G.; Labhardt, A. M.; Bax, A. *Biochemistry* **1992**, *31*, 8180–8190.

(49) LeMaster, D. M.; Richards, F. M. *Biochemistry* **1988**, *27*, 142–150.

(50) Arrowsmith, C. H.; Pachter, R.; Altman, R. B.; Iyer, S. B.; Jardetzky, O. *Biochemistry* **1990**, *29*, 6332–6341.